

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/501,930
Source: PCJ
Date Processed by STIC: 07/19/2005

ENTERED

BEST AVAILABLE COPY

20051



PCT

RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/501,930

TIME: 08:48:50

Input Set : A:\P25687.ST25.txt

Output Set : N:\CRF4\07192005\J501930.raw

```

3 <110> APPLICANT: TAKASHIMA, Shou
4     TSUJIMOTO, Masafumi
5     TSUJI, Shuichi
7 <120> TITLE OF INVENTION: GLYCOSYLATING ENZYME
9 <130> FILE REFERENCE: P25687
11 <140> CURRENT APPLICATION NUMBER: US 10/501,930
12 <141> CURRENT FILING DATE: 2004-07-29
14 <150> PRIOR APPLICATION NUMBER: PCT/JP03/00883
15 <151> PRIOR FILING DATE: 2003-01-30
17 <160> NUMBER OF SEQ ID NOS: 30
19 <170> SOFTWARE: PatentIn version 3.3
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 398
23 <212> TYPE: PRT
24 <213> ORGANISM: Mouse
26 <400> SEQUENCE: 1
28 Met Arg Ser Gly Gly Thr Leu Phe Ala Leu Ile Gly Ser Leu Met Leu
29 1      5      10      15
32 Leu Leu Leu Leu Arg Met Leu Trp Cys Pro Ala Asp Ala Pro Ala Arg
33      20      25      30
36 Ser Arg Leu Leu Met Glu Gly Ser Arg Glu Asp Thr Ser Gly Thr Ser
37      35      40      45
40 Ala Ala Leu Lys Thr Leu Trp Ser Pro Thr Thr Pro Val Pro Arg Thr
41      50      55      60
44 Arg Asn Ser Thr Tyr Leu Asp Glu Lys Thr Thr Gln Ile Thr Glu Lys
45 65      70      75      80
48 Cys Lys Asp Leu Gln Tyr Ser Leu Asn Ser Leu Ser Asn Lys Thr Arg
49      85      90      95
52 Arg Tyr Ser Glu Asp Asp Tyr Leu Gln Thr Ile Thr Asn Ile Gln Arg
53      100     105     110
56 Cys Pro Trp Asn Arg Gln Ala Glu Glu Tyr Asp Asn Phe Arg Ala Lys
57      115     120     125
60 Leu Ala Ser Cys Cys Asp Ala Ile Gln Asp Phe Val Val Ser Gln Asn
61      130     135     140
64 Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val Glu Ser Lys Lys
65 145     150     155     160
68 His Ile Pro Ile Arg Glu Asn Ile Phe His Met Phe Pro Val Ser Gln
69      165     170     175
72 Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val Val Gly Asn Gly
73      180     185     190
76 Gly Ile Leu Asn Lys Ser Leu Cys Gly Ala Glu Ile Asp Lys Ser Asp
77      195     200     205
80 Phe Val Phe Arg Cys Asn Leu Pro Pro Ile Thr Gly Ser Ala Ser Lys

```

RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/501,930

TIME: 08:48:50

Input Set : A:\P25687.ST25.txt

Output Set: N:\CRF4\07192005\J501930.raw

```

81      210      215      220
84 Asp Val Gly Ser Lys Thr Asn Leu Val Thr Val Asn Pro Ser Ile Ile
85 225      230      235      240
88 Thr Leu Lys Tyr Gln Asn Leu Lys Glu Lys Lys Ala Gln Phe Leu Glu
89      245      250      255
92 Asp Ile Ser Thr Tyr Gly Asp Ala Phe Leu Leu Leu Pro Ala Phe Ser
93      260      265      270
96 Tyr Arg Ala Asn Thr Gly Ile Ser Phe Lys Val Tyr Gln Thr Leu Lys
97      275      280      285
100 Glu Ser Lys Met Arg Gln Lys Val Leu Phe Phe His Pro Arg Tyr Leu
101      290      295      300
104 Arg His Leu Ala Leu Phe Trp Arg Thr Lys Gly Val Thr Ala Tyr Arg
105 305      310      315      320
108 Leu Ser Thr Gly Leu Met Ile Ala Ser Val Ala Val Glu Leu Cys Glu
109      325      330      335
112 Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys Thr Ile Glu Asp
113      340      345      350
116 Thr Pro Leu Ser His His Tyr Tyr Asp Asn Met Leu Pro Lys His Gly
117      355      360      365
120 Phe His Gln Met Pro Lys Glu Tyr Ser Gln Met Leu Gln Leu His Met
121      370      375      380
124 Arg Gly Ile Leu Lys Leu Gln Phe Ser Lys Cys Glu Thr Ala
125 385      390      395
128 <210> SEQ ID NO: 2
129 <211> LENGTH: 3166
130 <212> TYPE: DNA
131 <213> ORGANISM: Mouse
133 <400> SEQUENCE: 2
134 cggagcggcg agtcggtgcc gcccgggctg cgcttcgccc cggcagcttt ggcgggcagg 60
136 acgcccgtgg ctccaggatga gatcgggggg cagcgtgttc gccctcatag gcagcctgat 120
138 gctgctgctc ctccctgcgta tgctctggtg cccagccgac gcgcctgccc gctccaggct 180
140 gttgatggag ggaagcagag aggacaccag tggtagctca gctgcaactga agacactctg 240
142 gagcccgaca accccggtac caccgaccag gaacagcaca tatctggatg agaagacaac 300
144 ccaaataaca gagaaatgca aagatctgca atatagcttg aactctttat ctaacaaaac 360
146 gagacggtag tctgaggatg actacctcca gaccatcaca aacatacaga gatgccccatg 420
148 gaaccggcaa gcagaagaat atgacaattt tagagcaaaa ctggcttcct gttgcgatgc 480
150 cattcaagac ttctgtggtt cccagaacaa cactccagtg gggactaaca tgagctacga 540
152 ggtggaaagc aagaaacaca tccccattcg agagaacatt ttccacatgt ttccagtgtc 600
154 gcagcctttt gtggactatc cctataacca gtgtgcagtg gttggtaatg ggggaattct 660
156 caacaagtct ctctgcggag cagaaattga taaatctgac ttctgttca ggtgtaacct 720
158 ccccccaatc acagggagcg ctagttaaaga tgttgaagc aaaacaaatc ttgtgactgt 780
160 caatcccagc attataaccc tgaagtacca gaatttgaag gagaagaaag cacagttttt 840
162 ggaggacatc tccacctatg gagatgcatt cctcctcctg ccagcatttt cctatcgggc 900
164 caacacaggc atctctttta aagtctacca aacactcaaa gagtcaaaaa tgaggcaaaa 960
166 ggttctcttc ttccatccca ggtacctgag acacctcgct cttttctgga gaactaaagg 1020
168 ggtgactgca taccgcttgt ccacaggctt gatgattgca agtgctgctg tgggaactgtg 1080
170 tgaaaacgtg aagctctacg gattctggcc tttctctaag actatcgaag acaccccact 1140
172 cagtcaccac tactatgata acatgttacc taagcatggt ttccaccaga tgcctaaaga 1200
174 atacagccaa atgctccagc tccatatgag aggaatcctc aaactgcaat tcagcaaatg 1260

```

RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/501,930

TIME: 08:48:50

Input Set : A:\P25687.ST25.txt

Output Set: N:\CRF4\07192005\J501930.raw

```

176 tgaacggct taacgtttct tagaaggaga ataatttcag gaggtggagt ggatgtgtca 1320
178 cagcatctcc aaaaagccaa tagaagaagg cacagagaaa gcatgaatta caaaggcgct 1380
180 ctcccacttg tctagaccaa agccaccgcg cccactcac tttgcagcct ccacgagtca 1440
182 ctcatctctca ccttcaacgt tctttctctg agaatagaga ccaaaacatc agacttggat 1500
184 aagtaaaatg agataathtt tcaaatacat atagaatttg atttgagcca ggggtctctca 1560
186 gaatgcttcc ttgttcctat ccatgatagc cattcccacc tttatcagag tggtaatgaa 1620
188 actgtgcaat tgtgccaaag accctttctg aagagaatgt ctgaatcatg cgccgagttt 1680
190 ttacacacag ctcttccttt ataaataaat ccttcccatt ctccctccta gtagagtaca 1740
192 gaaacaaaat acccttgatg attcaggaag aaaagtcttt tttacttagc aatgtgcctg 1800
194 cttctgattc agttcgcttg tgacattaag ctgggttggg gttttggttg gatttggggc 1860
196 gtttcttcac ttcttttgc tttttttcc ttacctttat cagtttgtat tcgagcttcc 1920
198 tgccttggga ttctgcaatt ctctctccca ctgacaggat caactcaatg acataaagta 1980
200 gttcaaacat ccattgcttc tcacatgttt tatccataaa gttactcatc tgattttatt 2040
202 taaaatagtg aacatctact tgatatcaga cccgaggacc atcctccatt ggagaatatg 2100
204 aagatattgt cactggcaga aaagcagggtg tgtgccatta attgataaga taccacaagc 2160
206 atcatcatgc cagttatgaa cacagtgtcg aaaggatcat agacaggggt gggttaaact 2220
208 gatcccagta gaataaactt cagtgtacct atttcaggga agagttaatt tcacaattaa 2280
210 aactagtaaa tgaaccaatt cttaggcaca ttaagtggat tctgagtaaa agaaagggaa 2340
212 cagcaggaga aagctgttcg cttggttctg attacccaaa tgagcatgct ggaaggagggt 2400
214 tgtgaggcta cgctaaaacc tctgcgtagg gagagagtac agtgcagtag tgtggcggct 2460
216 tttgtccaca ctggtgaagg gtgagtaatt cagagccaat cacatcaca ggatggacac 2520
218 acctaactca tcacttcagg gggagatgaa tgctttcatg agaaattaca ctcataagct 2580
220 aagcatcagt tttgagtaaa atttgagtag atgttaaata tgaacatttt atacctctta 2640
222 ctaatgtccc accgacacct tttaatgtaa gcacatttat ttattaagtt acttgacatt 2700
224 aaatgcttat gtctgtatat tctgttcac catcgatttt cccaaaaagt aagagcatag 2760
226 gagatgagga ctacatgcca agaaaactat aaattttact ctttaattct tacttgagcc 2820
228 agcttgttgt ttatcaagtg cttttttgaa gagacagcac cctgtgaatt cttcattctg 2880
230 atacagtgtc accttgtatt taacatttgt aatgttgttt caagtttaca tctctttcat 2940
232 tcttttatag caaatcaaac gtattagctt cagaaattta tcagaagttc atatataaat 3000
234 attttgcaaa gggtaaaagg cttttttgtt aaataaaata aaatttatta ttttcttctg 3060
236 atgaatagag gctcttttat gctgctgcta atgaacctaa ttagctttta attatctct 3120
238 agcaacattg gtcacgtttc aatcatgcta ttagcaaaaa aaaaaa 3166

```

241 <210> SEQ ID NO: 3

242 <211> LENGTH: 398

243 <212> TYPE: PRT

244 <213> ORGANISM: Human

246 <400> SEQUENCE: 3

```

248 Met Arg Pro Gly Gly Ala Leu Leu Ala Leu Leu Ala Ser Leu Leu Leu
249 1 5 10 15
252 Leu Leu Leu Leu Arg Leu Leu Trp Cys Pro Ala Asp Ala Pro Gly Arg
253 20 25 30
256 Ala Arg Ile Leu Val Glu Glu Ser Arg Glu Ala Thr His Gly Thr Pro
257 35 40 45
260 Ala Ala Leu Arg Thr Leu Arg Ser Pro Ala Thr Ala Val Pro Arg Ala
261 50 55 60
264 Thr Asn Ser Thr Tyr Leu Asn Glu Lys Ser Leu Gln Leu Thr Glu Lys
265 65 70 75 80
268 Cys Lys Asn Leu Gln Tyr Gly Ile Glu Ser Phe Ser Asn Lys Thr Lys
269 85 90 95

```

RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/501,930

TIME: 08:48:50

Input Set : A:\P25687.ST25.txt

Output Set: N:\CRF4\07192005\J501930.raw

```

272 Gly Tyr Ser Glu Asn Asp Tyr Leu Gln Ile Ile Thr Asp Ile Gln Ser
273      100      105      110
276 Cys Pro Trp Lys Arg Gln Ala Glu Glu Tyr Ala Asn Phe Arg Ala Lys
277      115      120      125
280 Leu Ala Ser Cys Cys Asp Ala Val Gln Asn Phe Val Val Ser Gln Asn
281      130      135      140
284 Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val Glu Ser Lys Lys
285 145      150      155      160
288 Glu Ile Pro Ile Lys Lys Asn Ile Phe His Met Phe Pro Val Ser Gln
289      165      170      175
292 Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val Val Gly Asn Gly
293      180      185      190
296 Gly Ile Leu Asn Lys Ser Leu Cys Gly Thr Glu Ile Asp Lys Ser Asp
297      195      200      205
300 Phe Val Phe Arg Cys Asn Leu Pro Pro Thr Thr Gly Asp Val Ser Lys
301      210      215      220
304 Asp Val Gly Ser Lys Thr Asn Leu Val Thr Ile Asn Pro Ser Ile Ile
305 225      230      235      240
308 Thr Leu Lys Tyr Gly Asn Leu Lys Glu Lys Lys Ala Leu Phe Leu Glu
309      245      250      255
312 Asp Ile Ala Thr Tyr Gly Asp Ala Phe Phe Phe Leu Pro Ala Phe Ser
313      260      265      270
316 Phe Arg Ala Asn Thr Gly Thr Ser Phe Lys Val Tyr Tyr Thr Leu Glu
317      275      280      285
320 Glu Ser Lys Ala Arg Gln Lys Val Leu Phe Phe His Pro Lys Tyr Leu
321      290      295      300
324 Lys Asp Leu Ala Leu Phe Trp Arg Thr Lys Gly Val Thr Ala Tyr Arg
325 305      310      315      320
328 Leu Ser Thr Gly Leu Met Ile Thr Ser Val Ala Val Glu Leu Cys Lys
329      325      330      335
332 Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys Thr Val Glu Asp
333      340      345      350
336 Ile Pro Val Ser His His Tyr Tyr Asp Asn Lys Leu Pro Lys His Gly
337      355      360      365
340 Phe His Gln Met Pro Lys Glu Tyr Ser Gln Ile Leu Gln Leu His Met
341      370      375      380
344 Lys Gly Ile Leu Lys Leu Gln Phe Ser Lys Cys Glu Val Ala
345 385      390      395

```

348 <210> SEQ ID NO: 4

349 <211> LENGTH: 1500

350 <212> TYPE: DNA

351 <213> ORGANISM: Human

353 <400> SEQUENCE: 4

```

354 ggtggcggcg gcggcgggag ccgcgagtcg gggccgcccg ggctgtgctt cgccccggca      60
356 gcagcgggtg cggcggcgcc tgtggctcag gatcgggccg gggggcgcac tgctcgccct      120
358 gctcgccagc ctgctgctgc tgcctgctgc gcgctgctc tggcgcccg cagacgcgcc      180
360 cggccgcgcc aggtattctg tggaggaaag cagggaggcc acccagggca ccccgcgagc      240
362 gctgaggagc ctccggagcc cggcgaccgc ggtaccgcgc gccactaaca gcacatatct      300
364 gaatgagaag tcgctccaac tgacggagaa atgcaaaaat ctgcaatatg gcattgagtc      360

```

RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/501,930

TIME: 08:48:50

Input Set : A:\P25687.ST25.txt

Output Set: N:\CRF4\07192005\J501930.raw

```

366 tttctctaac aaaacgaaag ggtattcaga gaacgactac cttcagatta tcacagatat 420
368 acagagttgt ccatggaaac ggcaagcaga agaatatgca aatttttagag ccaaacttgc 480
370 ttcctgctgt gatgctgttc aaaactttgt tgtttctcag aataacactc cagttgggac 540
372 taatatgagt tacgaggtgg aaagcaaaaa agaaatccca attaagaaga acatttttca 600
374 tatgtttcca gtgtcccagc cttttgtgga ctacccttat aatcagtggt cagtggtcgg 660
376 aaatggggga attctgaata agtctctctg tggaaactgaa atagataaat ccgacttcgt 720
378 ttttaggtgt aacctacccc caaccacagg agatgttagt aaagatgttg gcagtaaaac 780
380 aaatccttggt actataaaatc caagcatcat aactctgaaa tatgggaact taaaggaaaa 840
382 aaaagcccta tttttggagg acattgcaac ctatggagat gcattttttt ttctgccagc 900
384 attttccttc agggccaaca cgggtacctc tttcaaagta tactacacgc tcgaagagtc 960
386 taaagcaaga caaaagggttc tatttttcca tcccaagtac ctgaaagatc tggccctttt 1020
388 ctggagaact aaaggtgtga ctgcataccg cttgtccacc ggcttgatga tcacaagtgt 1080
390 tgcagtggaa ctgtgtaaaa atgtgaagct gtatggattc tggcccttct ctaaaactgt 1140
392 agaagacata cctgtcagcc atcactatta tgacaacaag ctacctaaac atgggtttcca 1200
394 tcagatgccc aaagaataca gccagatcct ccaacttcac atgaaaggaa tcctcaaact 1260
396 gcaatttagc aaatgtgaag tcgcctaaac aaagtatctt aaaatgggaa taattttaat 1320
398 ataatgcagt aggtgattaa caatgtctcc aaacacaaaa ggaggtggct aaagagtatt 1380
400 ttgagatgag ccccaaaatt tggtttgacc aaagcttccc cactcatttt gcaatgatgg 1440
402 caagtcattc aatccttctc atcttcattt tttctcctta taacatggac accatatctg 1500

```

405 <210> SEQ ID NO: 5

406 <211> LENGTH: 529

407 <212> TYPE: PRT

408 <213> ORGANISM: Human

410 <400> SEQUENCE: 5

```

412 Met Lys Pro His Leu Lys Gln Trp Arg Gln Arg Met Leu Phe Gly Ile
413 1 5 10 15
416 Phe Ala Trp Gly Leu Leu Phe Leu Leu Ile Phe Ile Tyr Phe Thr Asp
417 20 25 30
420 Ser Asn Pro Ala Glu Pro Val Pro Ser Ser Leu Ser Phe Leu Glu Thr
421 35 40 45
424 Arg Arg Leu Leu Pro Val Gln Gly Lys Gln Arg Ala Ile Met Gly Ala
425 50 55 60
428 Ala His Glu Pro Ser Pro Pro Gly Gly Leu Asp Ala Arg Gln Ala Leu
429 65 70 75 80
432 Pro Arg Ala His Pro Ala Gly Ser Phe His Ala Gly Pro Gly Asp Leu
433 85 90 95
436 Gln Lys Trp Ala Gln Ser Gln Asp Gly Phe Glu His Lys Glu Phe Phe
437 100 105 110
440 Ser Ser Gln Val Gly Arg Lys Ser Gln Ser Ala Phe Tyr Pro Glu Asp
441 115 120 125
444 Asp Asp Tyr Phe Phe Ala Ala Gly Gln Pro Gly Trp His Ser His Thr
445 130 135 140
448 Gln Gly Thr Leu Gly Phe Pro Ser Pro Gly Glu Pro Gly Pro Arg Glu
449 145 150 155 160
452 Gly Ala Phe Pro Ala Ala Gln Val Gln Arg Arg Arg Val Lys Lys Arg
453 165 170 175
456 His Arg Arg Gln Arg Arg Ser His Val Leu Glu Glu Gly Asp Asp Gly
457 180 185 190
460 Asp Arg Leu Tyr Ser Ser Met Ser Arg Ala Phe Leu Tyr Arg Leu Trp

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/501,930

DATE: 07/19/2005

TIME: 08:48:52

Input Set : A:\P25687.ST25.txt

Output Set: N:\CRF4\07192005\J501930.raw